

# Temporal synchrony is an effective cue for grouping and segmentation in the absence of form cues

Reuben Rideaux

Research School of Psychology,  
The Australian National University, Canberra,  
Australian Capital Territory, Australia



David R. Badcock

School of Psychology,  
The University of Western Australia, Crawley,  
Western Australia, Australia



Alan Johnston

School of Psychology, The University of Nottingham,  
Nottingham, United Kingdom



Mark Edwards

Research School of Psychology,  
The Australian National University, Canberra,  
Australian Capital Territory, Australia



The synchronous change of a feature across multiple discrete elements, i.e., temporal synchrony, has been shown to be a powerful cue for grouping and segmentation. This has been demonstrated with both static and dynamic stimuli for a range of tasks. However, in addition to temporal synchrony, stimuli in previous research have included other cues which can also facilitate grouping and segmentation, such as good continuation and coherent spatial configuration. To evaluate the effectiveness of temporal synchrony for grouping and segmentation in isolation, here we measure signal detection thresholds using a global-Gabor stimulus in the presence/absence of a synchronous event. We also examine the impact of the spatial proximity of the to-be-grouped elements on the effectiveness of temporal synchrony, and the duration for which elements are bound together following a synchronous event in the absence of further segmentation cues. The results show that temporal synchrony (in isolation) is an effective cue for grouping local elements together to extract a global signal. Further, we find that the effectiveness of temporal synchrony as a cue for segmentation is modulated by the spatial proximity of signal elements. Finally, we demonstrate that following a synchronous event, elements are perceptually bound together for an average duration of 200 ms.

## Introduction

The need to detect objects moving within the visual field is a common visual task. This is achieved with

ease; however, the apparent effortlessness of this feat undermines the complexity of the process. Receptive fields in the early visual system are relatively small, so the objects that we view typically stimulate the receptive fields of many neurons (Hubel & Wiesel, 1962, 1965). To fully process these objects, the output of the corresponding population of neurons needs to be combined; this holds for both form and motion processing and is achieved through segmentation and pooling. Segmentation goes hand-in-hand with pooling; that is, while signals which belong to a common object need to be pooled together, they must also be segmented from those which belong to other objects. The desired outcome of these processes is to increase the signal to noise ratio of the neural processing linked to the perception of these objects.

Visual regularities within the environment are detected by the visual system and used to indicate the boundaries of objects, allowing common elements of an object to be bound together—for example, visual grouping—while distinguishing it from other objects and the background—for example, figure-ground segmentation (Driver & Baylis, 1996; Kovács & Julesz, 1993; Nakayama, Shimojo, & Silverman, 1989). Grouping and segmentation cues can be both spatial and temporal in nature. A potential class of temporal cue is temporal synchrony; that is, elements that change an aspect of how they are defined at the same time may be bound together and segmented from other signals.

Citation: Rideaux, R., Badcock, D. R., Johnston, A., & Edwards, M. (2016). Temporal synchrony is an effective cue for grouping and segmentation in the absence of form cues. *Journal of Vision*, 16(11):23, 1–12, doi:10.1167/16.11.23.

doi: 10.1167/16.11.23

Received May 22, 2016; published September 30, 2016

ISSN 1534-7362



This is particularly relevant for motion signals, given their dynamic nature. However, even in form processing, evidence indicates that temporal synchrony can facilitate contour integration (Bex, Simmers, & Dakin, 2001; Lee & Blake, 2001; Usher & Donnelly, 1998) and detection of apparent spatial location/form (Kandil & Fahle, 2004; Lee & Blake, 1999). There is even evidence suggesting that temporal synchrony can facilitate segmentation across modalities (Kösem & van Wassenhove, 2012). In regards to visual perception, temporal cues appear to help most when spatial cues are ambiguous or absent in a static version of the display, although care must be taken to avoid spatial cues arising from low pass temporal filtering (Adelson & Farid, 1999; Farid, 2002; Farid & Adelson, 2001; for a review, see Blake & Lee, 2005).

With motion stimuli, a temporal cue that was noted by the Gestalt psychologists is common-fate, i.e., objects that move in the same direction are bound together. However, when not combined with form cues, common-fate doesn't appear to be particularly strong. For example, in the original global-motion stimulus developed by Newsome and Pare (1988), the signal dots cannot be fully segmented from the noise dots, as demonstrated by the linear tuning of V5/MT cells to signal coherence level (Britten, Shadlen, Newsome, & Movshon, 1993; see also Edwards & Badcock, 1998) and that adding form information to the arrangement of the signal dots dramatically lowers global-motion thresholds (Edwards, 2009).

With motion stimuli, like with static form stimuli, temporal synchrony can also be generated by changing the properties of a subgroup of the elements in unison, e.g., the contrast of the elements. A number of studies have investigated the role of temporal synchrony in pooling and segmentation; however, they have all, to varying degrees, also included form cues in the stimuli. Note the use of "form cues" here refers to texture and other spatial cues as well. For example, Alais, Blake, and Lee (1998) demonstrated the role of temporal synchrony using a display in which the motion of four separate components can either be perceived as a single global diamond shape moving upwards, or as four discrete local Gabors. The local motion cues are compatible with both local and global solutions, resulting in a bistable percept. The authors modulated the contrast of each component and showed that by manipulating whether modulation occurred asynchronously or synchronously across the components, perception could be biased towards either the local or global solution, respectively. However, their stimulus contained form information, i.e., an unambiguous diamond shape, which has been shown to augment temporal synchrony-based grouping (Leonards, Singer, & Fahle, 1996; Tang, Dickinson, Visser, Edwards, & Badcock, 2015; Usher & Donnelly, 1998) in excess of

the probability summation of the two properties (Lee & Blake, 2001). Similarly, Lee and Blake (1999, 2001) and Morgan and Castet (2002) demonstrated the capacity of temporal synchrony to drive segmentation using static stimuli. However, in their stimuli the signal and noise elements were spatially segregated, acting as a form cue which could be used to facilitate segmentation, as noted by Adelson and Farid (1999).

A number of studies have shown that, with static stimuli, form information tends to dominate temporal synchrony as a segmentation cue. Kiper, Gegenfurtner, and Movshon (1996) investigated the ability of observers to detect the spatial arrangement of a texture defined region made up of oriented bars. The authors flickered the elements that made up the target region, such that all the elements were either in phase, and different to the background, or the phase of all the elements (signal and noise) was randomized. However, this appeared to have no effect on grouping, as measured by the ability to detect the orientation of the region, beyond that facilitated by the orientation cues that were present. More recently, Lorenceau and Lalanne (2008) used outlines of geometrical shapes moving behind apertures that concealed their vertices such that determining the global solution required pooling of the correct subset of local motion components. While they found that form properties such as good continuation and spatial configuration of local components influenced segmentation, their results suggested that synchronous temporal modulation, i.e., flickering, has little effect on selection.

Given these findings, and the consistent presence of form cues in previous studies investigating this issue, here we are concerned with the ability to use temporal synchrony as a cue in the absence of form cues. To investigate this we employed a global-Gabor stimulus (Amano, Edwards, Badcock, & Nishida, 2009), which is a field of Gabor elements given random orientations, and speeds of carrier motion consistent with a globally coherent solution are applied to those elements carrying the signal (see Figure 1). When the otherwise identical Gabor elements are arranged with random orientations, the stimulus contains no salient form cues. Extracting the global motion signal from the stimulus requires observers to pool information across multiple elements to determine an intersection of constraints (IOC) solution, i.e., a rigid motion solution that would be consistent with the set of speed and direction combinations present in the array of Gabors (Adelson & Movshon, 1982). Another advantage of this stimulus is that it provides a sensitive metric to determine the effectiveness of the temporal cue: the signal-to-noise ratio required to determine the global motion direction. The stimulus is similar to that used by Lee and Blake (1999, 2001) and Morgan and Castet (2002), but they were not examining the ability to group for IOC

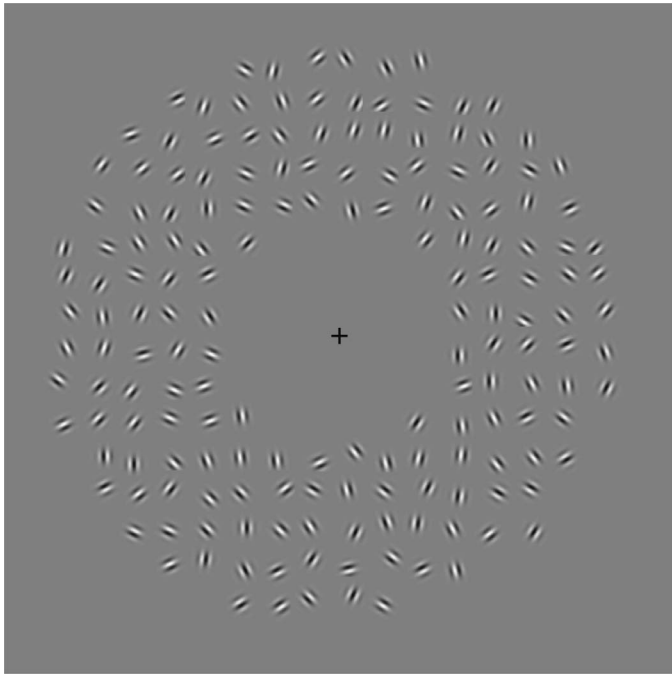


Figure 1. An example of the stimulus used in Experiment 1. The orientation of each Gabor is selected at random on each trial, and its corresponding drift rate is consistent with either a constant 2D IOC solution (signal) or a random solution between  $1^{\circ}$ – $360^{\circ}$  (noise).

pooling; and our signal elements are randomly intermingled among the noise elements, thus avoiding the potential confounds noted by Adelson and Farid (1999).

## Experiment 1: Temporal synchrony as a cue for grouping local motion elements

The aim of Experiment 1 was to determine the effectiveness of temporal synchrony as a grouping cue for motion pooling, independent of form cues. Specifically, we examined the strength of temporal synchrony as a cue for grouping a subset of one-dimensional (1D) local-motion signals in a global-Gabor stimulus.

## Method

### Participants

Fifteen observers participated in the experiment (mean age, 22). All had normal or corrected-to-normal visual acuity and gave informed written consent to participate in the study. All observers were naïve

regarding the aims of the study and were either given research credit or compensated \$15 for participation.

### Apparatus

All experiments were run under the MATLAB (version R2013a) programming environment, using software from PsychToolbox (Brainard, 1997; Pelli, 1997). Stimuli were presented on a Phillips Brilliance 202P4 CRT monitor that was driven by an Intel Iris graphics card in a host MacBook Pro computer. The monitor had a display size of  $406 \times 305$  mm, spatial resolution of  $1024 \times 768$  pixels, and a frame rate of 120 Hz.

### Stimuli

The stimuli were a modified version of the global-Gabor stimuli (Amano et al., 2009). It consisted of 172 Gabors, evenly positioned with regular separation between elements, placed within an annulus centered on fixation (inner ring radius,  $4^{\circ}$  visual angle; outer ring radius,  $11^{\circ}$ ), but with no lines defining the borders. Each Gabor (contrast, 0.5; spatial frequency, 3 cycles/ $^{\circ}$ ; phase,  $0^{\circ}$ ) was positioned within a Gaussian envelope (radius,  $0.66^{\circ}$ ; *SD*,  $0.16^{\circ}$ ), and its orientation was selected randomly. The drift rate of signal elements was consistent with a constant two-dimensional (2D) motion direction; that is, the speed at which signal elements drifted varied as a function of their orientation to conform to a global IOC solution. The global speed was set at  $2.22^{\circ}/s$ ; thus, the local drift rate of signal elements ranged from  $0^{\circ}$  (elements parallel to the motion direction) to  $2.22^{\circ}/s$  (elements perpendicular to the motion direction). In contrast, the drift rate of each noise element ( $0^{\circ}$ – $2.22^{\circ}/s$ ) was consistent with a different 2D direction ( $1^{\circ}$ – $360^{\circ}$ ) selected randomly, resulting in an incoherent global IOC solution when pooled. The background was gray (mean luminance,  $12 \text{ cd}/\text{m}^2$ ). An example frame of the stimulus is shown in Figure 1.

A single interval forced-choice (SIFC) procedure was employed; the observers' task was to indicate the direction of the global IOC solution. On each trial the signal direction ( $45^{\circ}$  or  $135^{\circ}$ , clockwise from  $0^{\circ}$  as up) was selected at random. The aim of the experiment was to examine whether temporal synchrony could be used to segment signal from noise elements in order to extract an IOC solution. Thus, we removed Gabors oriented around horizontal ( $90^{\circ}$ ) that would drift up or down depending on the signal direction, as these were easier to find in the field and could potentially be used alone to solve the task. Removing these orientations reduced this risk and encouraged observers to employ a strategy of pooling information from multiple signal elements. Thus, a range of possible Gabor orientations



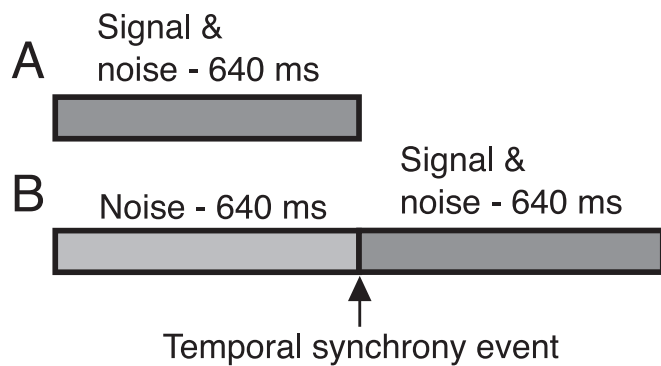


Figure 2. A schematic of the time-course of the presentation content (i.e., signal and/or noise) in Experiment 1. The (A) standard global motion time condition consists of a 640 ms presentation of both signal and noise elements, while the (B) other four conditions (signal phase shift, contrast spike, common-fate, and signal-noise phase shift) have an additional 640 ms presentation of noise elements preceding this. The defining feature of these four conditions is the temporal synchrony event (or lack of) which occurs at the onset of the signal and noise segment.

between  $0^{\circ}$ – $65^{\circ}$  and  $105^{\circ}$ – $180^{\circ}$  was used. In addition to this measure of prevention, a control experiment—described later—was run to further rule out the possibility that observers were extracting the global direction from the drift rate of a single element.

As a baseline from which other conditions could be compared, a standard global motion condition was run (Amano et al., 2009). That is, while the carrier motion of all signals remained constant for the duration of the presentation, a proportion of the elements moved in a coherent direction (signal) and the carrier motion for the remaining elements was chosen at random (noise). In this condition the stimulus was presented for 640 ms. The four other conditions contained different synchrony cues. In these conditions, the stimulus was presented for 1280 ms, with the first half (640 ms) of the presentation consisting of random carrier motion in each element (100% noise). The drift rate of the signal elements then changed to conform to the global IOC solution for the second half (640 ms) of the presentation. Thus, the second half of the presentation was identical to that used in the standard global motion condition. The defining feature of the four additional conditions was the event that occurred at the onset of the second (combined signal and noise) half of the presentation. A schematic depicting the time-course of the presentation content is shown in Figure 2.

To examine the effectiveness of temporal synchrony as a segmentation cue, temporal synchrony of the signal elements was induced in two ways: In one condition, all signal elements were “phase shifted” by  $180^{\circ}$  at the onset of the second interval (resulting in a pulse-like percept); in the other condition, the contrast

of the signal elements was increased for one frame (32 ms) to 70% before returning to the original level (50%). These conditions will be referred to as signal phase shift and contrast spike, respectively. Note that in these conditions, in addition to the temporal synchrony event, a form of common-fate cue is also present due to the coherent motion drift of the signal elements following the event. However, the magnitude of this common-fate varied between elements, depending on their orientation. That is, the gradient of potential for common-fate grouping of elements ranged from the strongest for those oriented perpendicular to the target motion direction to the weakest for those oriented in parallel. It should also be noted that this form of common-fate is different from the standard type, which refers to grouping of 2D motion signals, as here the signals are 1D and are common only at the IOC global-processing level. To evaluate the contribution of this form of common-fate as a segmentation cue, a (common-fate) condition was run identical to the other temporal synchrony conditions, but without either the signal phase shift or the contrast spike manipulations.

In the aforementioned temporal synchrony conditions, this grouping property is manipulated to determine whether it can facilitate segmentation of signal from noise elements. Another way to test the strength of segmentation would be to examine the detrimental effect on motion pooling resulting from grouping noise and signal elements together. Thus, a final condition was run to examine whether temporal synchrony would impede segmentation by grouping all elements together. This was achieved using the same stimulus as in the signal phase shift condition, with the exception that all elements were phase shifted. This condition will be referred to as signal + noise phase shift. However, note that common-fate cues were still present in this condition and may continue to facilitate segmentation beneficial to performing the task. That is, after the synchronous phase shift on all the elements, the signal Gabors still underwent the common-fate change.

### Procedure

The observers sat 50 cm from the monitor, with their head supported on a chin rest, and used the “z” and “/” keyboard keys to indicate the direction of the stimulus. An adaptive staircase procedure was employed using software from the Palamedes Toolbox (Prins & Kingdom, 2009), varying the ratio of signal-to-noise elements. The staircase uses a “psi-marginal” adaptive method, based on Kontsevich and Tyler’s (1999) psi-method (Prins, 2013). For each condition 50 trials were run using the adaptive staircase, with the order of conditions randomized between observers.

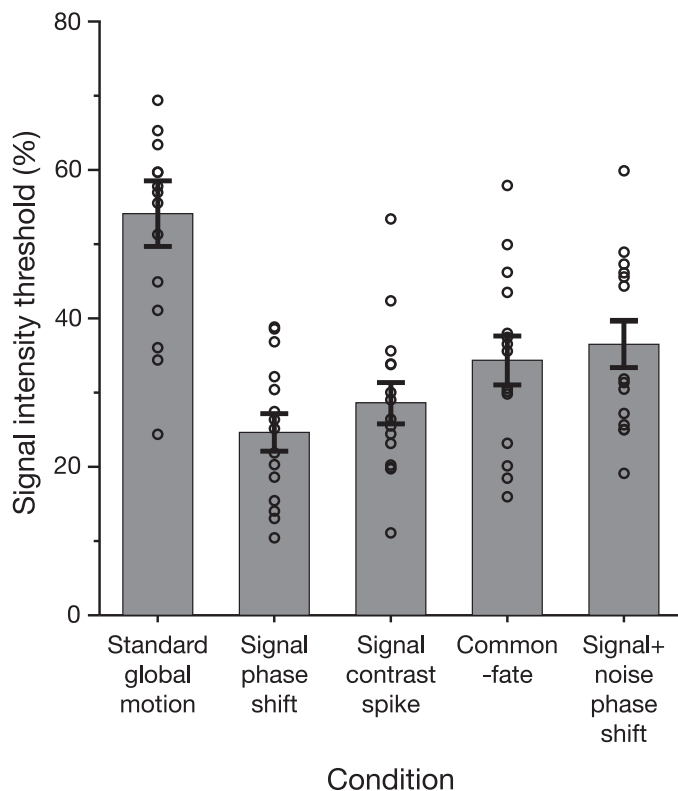


Figure 3. The average signal intensity thresholds across observers for each of the conditions in Experiment 1. The black circles indicate individual data points. Error bars represent  $\pm 1$  SEM.

## Results

The average signal intensity threshold across observers for each condition is shown in Figure 3. Note that the average threshold in the standard global motion condition, which has comparable parameters to that used by Amano et al. (2009), is considerably higher than previous estimates ( $\sim 15\%$ ). However, this may be due to a combination of employing naïve observers and omitting a proportion of carrier orientations during construction of the stimulus in the current experiment.

A one-way, repeated-measures analysis of variance (ANOVA) revealed a significant effect of condition,  $F(4, 56) = 16.6$ ,  $p < 0.001$ . Posthoc comparisons were run to evaluate differences between selected conditions, using a Bonferroni correction to adjust for multiple comparisons. There was no difference between the signal phase shift and contrast spike conditions,  $p = 1.0$ , 95% CIs  $[-14.8; 6.8]$ , while the average thresholds were lower in these conditions than in the standard global motion condition,  $ps < 0.001$ , CIs  $[-14.8, -44.1]$ ,  $[-11.3, -39.7]$ , respectively. The mean threshold in the common-fate condition was lower than in the standard global motion condition,  $p = 0.04$ , CIs  $[-38.7, -0.7]$ , while also higher than in the signal phase shift condition,  $p = 0.02$ , CIs  $[1.1, 18.4]$ , but not the contrast

spike condition,  $p = 1.0$ , CIs  $[-6.1, 17.7]$ . This indicates that, perhaps unsurprisingly, even this form of common-fate is an effective segmentation cue; however, it also demonstrates that it cannot (alone) account for the improvement in performance observed in the signal phase shift condition. In other words, the strength of the common-fate cue was not as robust as the (phase shift) temporal synchrony cue. These results also suggest that the phase shift may be stronger cue for segmentation than the contrast spike.

The average threshold in the signal + noise phase shift condition was lower than in the standard global motion condition,  $p = 0.02$ , CIs  $[-32.8, -2.4]$ , indicating that attempting to group both signal and noise elements together with a  $180^\circ$  phase shift did not have a deleterious effect on performance, as we had anticipated. The mean threshold in this condition was the same as in the common-fate condition,  $p = 1.0$ , CIs  $[-10.4, 14.6]$ ; thus, it is likely that phase shifting all the elements had no additional effect as there was no signal-versus-noise distinction, so the difference in threshold between the signal + noise phase shift and standard global motion conditions simply reflects the influence of common-fate segmentation. Interestingly, this suggests that a large field contrast reversal does not mask a subsequent motion onset.

We are confident that exclusion of orientations around  $90^\circ$  during the construction of the stimulus was sufficient to prevent the task from being performed without pooling across a number of signal elements. Indeed, in the signal phase shift and contrast spike conditions, the average threshold number of elements required to perform the task ranged from 40–50, showing that the presence of a considerable number of signal elements was necessary to extract the global 2D solution. However, given the importance of distinguishing between observers using temporal synchrony as a cue to segment signal from noise elements and using it to selectively attend to a single element, in order to perform the task, we ran a control experiment to further scrutinize this possibility.

## Control experiment

The control experiment consisted of two conditions which were the same as the standard global motion and signal phase shift conditions used in Experiment 1, with the exception that instead of two possible global directions, eight possible directions (cardinals and obliques) were employed, and no orientations were excluded. Thus, determining the correct direction, from the eight alternative choices, was virtually impossible without pooling information from multiple signal elements. The same 15 observers from the main experiment participated in the control; their average

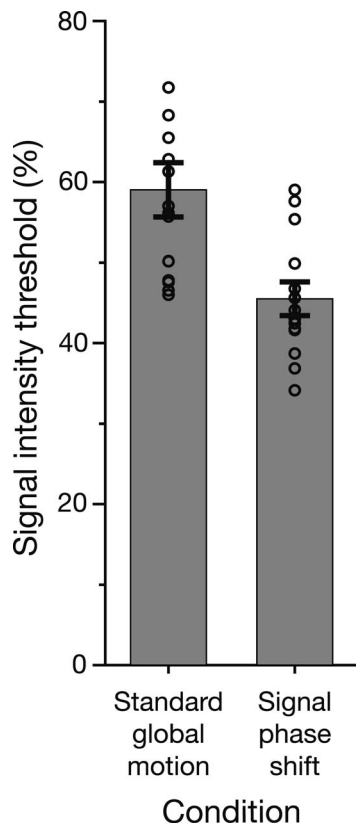


Figure 4. The average signal intensity thresholds across observers for the control conditions in Experiment 1. The black circles indicate individual data points. Error bars represent  $\pm 1$  SEM.

signal intensity thresholds for the two control conditions are shown in Figure 4.

A paired  $t$  test revealed that the average signal intensity threshold in the signal phase shift control condition was lower than that in the standard global motion condition,  $t(14) = 4.3$ ,  $p < 0.001$ . This supports the interpretation that a grouping/segmentation mechanism improved performance in the main conditions (signal phase shift and contrast spike) of Experiment 1, rather than facilitating selective attention to a single Gabor to resolve the direction.

## Discussion

The results of Experiment 1 show that temporal synchrony is a powerful cue for signal-versus-noise segmentation of local motion components. Furthermore, they demonstrate that it is effective even in the absence of form cues. While IOC-based common-fate can also be used to group elements together, the results demonstrate that other forms of temporal synchrony improved the segmentation of elements beyond that achievable through common-fate alone when the spatial configuration of the elements is randomized.

In addition to temporal and form cues for pooling and segmentation, another strong cue that has been investigated is spatial proximity. Forte, Hogben, and Ross (1999) and Motoyoshi (2004) demonstrated a positive relationship between the spatial proximity of elements and observers' sensitivity for detecting asynchronous activity. That is, the closer signal elements were spatially positioned, the more easily observers were able to detect temporal asynchronies between them. However, again, a potential artifact in these studies may have been the present due to the spatial segregation of signal elements (Adelson & Farid, 1999). In contrast, the spatial proximity of signal elements within a mixture of signal and noise elements has no impact on standard global form (Dickinson, Broderick, & Badcock, 2009) or motion processing (Morley & Badcock, 2016), unless the observer has prior knowledge of the location of the subset of signal elements, or signal intensity is at supra threshold levels (Greenwood & Edwards, 2009). Thus, if spatial proximity operates to augment the grouping mechanisms of temporal synchrony, one would expect that when signal elements are arranged in high spatial proximity, fewer elements will be required to extract a global 2D solution, but only in the presence of a synchronous event. Experiment 2 examines this possibility by comparing performance between standard global motion processing and motion processing facilitated by temporal synchrony, when signal elements are more/less proximal.

## Experiment 2: Temporal synchrony and spatial proximity

Previous research suggests that the spatial proximity of elements influences the ability of observers to detect asynchronous activity (Forte, Hogben, & Ross, 1999; Motoyoshi, 2004), indicating that spatial proximity influences the effectiveness of segmentation facilitated by temporal synchrony. However, due to the presence of other form cues in these studies, it is unclear whether spatial proximity impacts temporal synchrony in the absence of form cues. Furthermore, given that the spatial proximity of signal components does not influence standard global motion processing at threshold levels, this condition provides an opportunity to distinguish the mechanisms engaged when a synchronous event is present versus absent, as distinct predictions can be made regarding the effect of manipulating the spatial proximity of signal elements when driving standard global motion compared to motion processing facilitated by temporal synchrony. That is, if increased spatial proximity of signal elements improves the effectiveness of temporal synchrony in the absence of other form cues, by increasing the spatial proximity

signal elements in the standard global motion and signal phase shift (henceforth referred to as phase shift) conditions employed in Experiment 1, no change in performance should be observed in the standard global motion condition, while signal intensity thresholds in the phase shift condition should be reduced. Here we assess this possibility, examining the impact of spatial proximity on standard global motion processing and motion processing facilitated by temporal synchrony, in the absence of additional form cues.

## Method

### Participants

The same fifteen observers participated in the experiment (mean age, 22). All had normal or corrected-to-normal acuity and gave informed written consent to participate in the study. All observers were naïve regarding the aims of the study and were either given research credit or compensated \$10 for participation.

### Stimuli and procedure

To contrast the impact of the spatial proximity of signal elements during standard global motion processing with motion processing when elements are bound together through temporal synchrony, here we employed the same stimuli and procedure as in the standard global motion and phase shift conditions of Experiment 1, while manipulating the spatial proximity of signal elements. Two spatial proximity conditions were employed, one (low spatial proximity) condition where signal positioning was the same as in Experiment 1, and one (high spatial proximity) condition where the random allocation of signal elements was restricted to one half, i.e., top, bottom, left, or right, of the annulus. The half of the annulus where the signal elements were positioned was selected at random on each trial. If the number of signal elements exceeded the capacity of half the annulus, i.e., signal intensity  $> 50\%$ , they were randomly allocated throughout the next adjacent rows/columns. Thus, a  $2 \times 2$  experimental design was employed: segmentation (standard global motion/phase shift)  $\times$  spatial proximity (low/high).

## Results and discussion

A two-way, repeated-measured ANOVA was used to compare performance between the four conditions (shown in Figure 5). There were main effects of segmentation and spatial proximity,  $F(1, 14) = 5.7$ ,  $p = 0.03$  and  $F(1, 14) = 28.2$ ,  $p < 0.001$ , respectively; however, there was no interaction,  $F(1, 14) = 1.0$ ,  $p = 0.32$ . Paired  $t$  tests revealed that there was no difference

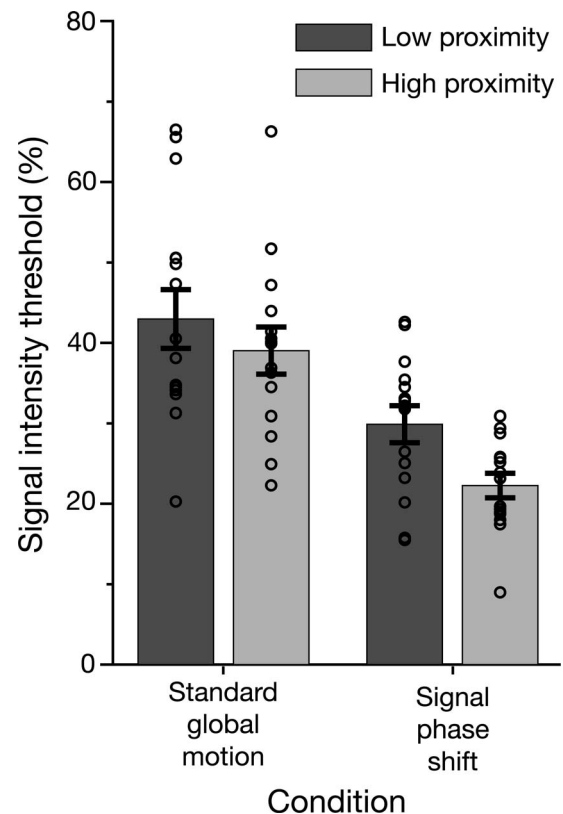


Figure 5. The average signal intensity thresholds across observers for each of the conditions in Experiment 2. The black circles indicate individual data points. Error bars represent  $\pm 1$  SEM.

between low and high spatial proximity for standard global motion conditions,  $t(14) = 1.1$ ,  $p = 0.28$ . In contrast, within the phase shift conditions, the average threshold was significantly lower in the high spatial proximity condition,  $t(14) = 3.1$ ,  $p = 0.007$ .

These results show that increased spatial proximity enhances the capacity of temporal synchrony to act as a cue for segmentation. Furthermore, by demonstrating that spatial proximity does not have an impact on performance in the standard global motion condition, this outcome clearly shows a distinction between—at least some of—the mechanisms engaged when signal elements are grouped through a synchronized event.

Experiment 2 examined the influence of spatial proximity on the effectiveness of temporal synchrony as a binding cue. Another aspect of grouping through temporal synchrony, which this stimulus can uniquely be used to examine, is temporal proximity. That is, after a synchronous event segments signal from noise, how long is that segmentation perceptually maintained without subsequent/additional grouping properties? Experiment 3 investigates this question by determining the duration for which a synchronous event continues to facilitate motion processing when no other grouping properties are present to maintain segmentation.



## Experiment 3: Temporal synchrony and temporal proximity

Having demonstrated the importance of spatial proximity for segmentation facilitated by temporal synchrony in Experiment 2, here we investigate the influence of proximity along the dimension of time, i.e., temporal proximity. Using the stimulus employed in Experiments 1 and 2, we temporally separate the synchronous event, i.e., the phase shift, and the point at which the elements grouped by this event begin to provide the information which can be used to perform the task, i.e., move in the target direction. Thus, by separating these events, we can investigate the duration for which elements are perceptually bound together through temporal synchrony.

### Method

#### Participants

Fifteen new observers participated in the experiment (mean age, 23). All had normal or corrected to normal acuity and gave informed written consent to participate in the study. All observers were naïve regarding the aims of the study and were either given research credit or compensated \$10 for participation.

#### Stimuli

The stimuli and procedure were similar to those used in the phase shift and common-fate conditions in Experiment 1. In Experiment 1, reduced detection thresholds in the phase shift, relative to the common-fate, condition indicated that local signal elements had been, at least partially, segmented from the noise elements as a result of the synchronous 180° phase shift. Here, to determine the persistence of this segmentation, in the absence of other grouping cues, we first established each observer's signal intensity threshold in the phase shift condition. We then ran a second tailored condition, where the signal intensity was fixed at the previously obtained threshold value, varying the duration between the synchronous event (phase shift) and the point where the signal elements began to drift in the signal direction (signal drift). In order to perform the task at this signal intensity, the elements would have to be perceptually grouped by the synchronous event; thus, as the temporal gap widened between the synchronous event and the signal direction drift change, eventually the segmentation produced by the synchronous event would decay and the observer could no longer be capable of performing the task.

Given the importance of demonstrating that the task could not be performed at the threshold employed in the

second condition without effective segmentation, a second (time two) phase shift and a common-fate condition were then run to determine if practice effects had resulted in reduced detection thresholds, thus allowing the task to be performed at the initial (time one) level obtained without additional segmentation cues.

### Results and discussion

The average threshold delay across all observers was 200 ms (range, 50–332 ms; 95% CIs, 150–251 ms) (Figure 6A). Paired  $t$  tests revealed that there was no difference between signal intensity thresholds at time one and two of the phase shift conditions,  $t(14) = 1.2$ ,  $p = 0.24$ , and thresholds in these conditions were both significantly lower than in the common-fate condition,  $t(14) = 3.4$ ,  $p = 0.004$  and  $t(14) = 2.6$ ,  $p = 0.02$ , respectively (Figure 6B).

Given that the signal intensity thresholds were the same between time one and two of the phase shift conditions, and significantly lower than in the common-fate condition, we are confident that the average threshold duration (200 ms) between the phase shift and signal drift reflects the average time course (of decay) of segmentation resulting from temporal synchrony, rather than being influenced by practice effects. This is further evidenced by observers' inability to perform the task at longer SOAs in the main condition, although the considerable interobserver variability in the time course should be noted.

The average rate of decay is within the bounds of iconic memory, which varies between 100–300 ms (Averbach & Coriell, 1961; Efron & Lee, 1971; Sperling, 1960) based on several factors including exposure duration, contrast, and form of information. However, this interpretation is unlikely for two reasons. First, for stimulus exposure durations of 130 ms or more, iconic memory only persists for approximately 100 ms (Efron, 1970). However, here the exposure duration of the stimulus preceding the synchronous event was 640 ms, yet effective segmentation was retained for ~ 200 ms. Second, the noise following the synchronous event would have operated as a backwards mask, immediately overriding any potential iconic persistence resulting from the phase shift.

A popular account, first proposed by Milner (1974) and further developed by von der Malsburg (1981/1994) and Crick and Koch (1990), describing the underlying neural mechanism of visual segmentation, claims that elements are bound together via synchronous activity across cell assemblies, with precision in the millisecond range. This activity reverberates within neural circuits and results in brief changes in synaptic efficiency between synchronized cells. Support for this hypothesis has been provided by both experimental studies and computational modelling (Abeles, Berg-



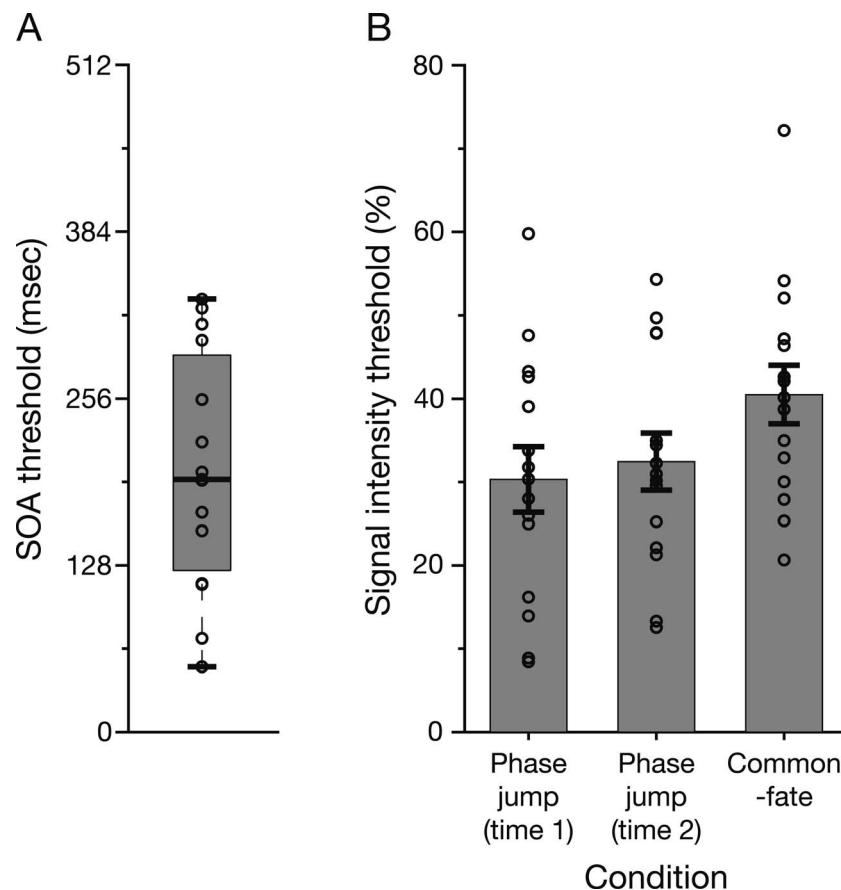


Figure 6. (A) A boxplot indicating the threshold stimulus onset asynchrony (SOA) between phase shift and signal drift at which observers could perform the task. (B) The average signal intensity thresholds across observers for each of the conditions in Experiment 3. Error bars represent  $\pm 1$  SEM. For both plots, the black circles indicate individual data points.

man, Margalit, & Vaadia, 1993; Diesmann, Gewaltig, & Aertsen, 1999). When considered in this framework, the decay rate of segmentation found here could reflect the decay (or echo) of this reverberation across cell assemblies, which without continuing visual segmentation cues, gradually becomes desynchronized.

## General discussion

The main findings of the current study are that temporal synchrony is an effective cue for segmentation in the absence of form cues (Experiment 1), the spatial proximity of signal elements augments the effectiveness of this segmentation (Experiment 2), and that segmentation decays after  $\sim 200$  ms in the absence of other grouping cues (Experiment 3). Whereas numerous studies have found that temporal synchrony can be used as a cue for grouping and segmentation, they have all—to varying degrees—contained additional form cues, making conclusions regarding the effectiveness of temporal synchrony in isolation speculative. Here we have demonstrated that temporal synchrony is an

effective cue for segmentation in isolation of additional form cues by randomly intermingling signal and noise elements. This is particularly important given previous studies which suggested this may not be the case (Adelson & Farid, 1999; Farid, 2002; Farid & Adelson, 2001; Kiper et al., 1996; Lorenceau & Lalanne, 2008).

Similarly, while previous research suggested that observers are more sensitive to asynchronous activity when elements are positioned closer together (Forte, Hogben, & Ross, 1999; Motoyoshi, 2004), the impact of spatial proximity in isolation could not be evaluated as the stimuli employed in these studies contained additional form information in the way of spatially segregated signal elements. Here we extend and clarify this previous work by demonstrating that spatial proximity moderates the effectiveness of temporal synchrony as a segmentation cue in isolation of additional form information.

In their review of temporal synchrony, Blake and Lee (2005) discuss two categories of temporal structure, deterministic and stochastic, arguing that the latter is more appropriate for the study of temporal synchrony for both ecological and practical justifications. Here we extend this by demonstrating the effect of a single temporal signal, as opposed to different temporal

structures, i.e., repetitive temporal spikes, either periodic or nonperiodic.

Arguably the most interesting finding of the current study is that relating to the time course of segmentation. The unique nature of the stimulus employed here allowed us to determine the duration which, without visual segmentation cues to maintain it, the benefits of binding elements together with a synchronous event are lost. To the best of our knowledge, this is the first demonstration of its kind; these results may provide useful insights into the neural mechanisms which initiate and maintain perceptual segmentation, e.g., neural synchrony. Goodbourn and Forte (2013) found evidence indicating that segmentation driven by temporal synchrony is achieved by neurons in the early stages of visual processing. The current paradigm could be used to extend this by manipulating the content of noise during the SOA between the synchronous event and the carrier signal, to examine precisely at what stage/s in the visual system the segmentation is maintained. Additionally, form cues could be manipulated, e.g., creating a hybrid paradigm of Experiments 2 and 3, to determine whether the decay rate is influenced by the initial strength of perceptual segmentation.

In summary, the present study clarifies and extends previous research on temporal synchrony, by demonstrating that it is an effective segmentation cue, even in the absence of form cues. We then go on to investigate both spatial and temporal proximity aspects of segmentation and demonstrate their respective influences on the effectiveness of this mechanism.

**Keywords:** temporal synchrony, intersection of constraints, IOC, motion processing, spatial temporal proximity

## Acknowledgments

This work was supported by an Australian Postgraduate Award to R.R. and Australian Research Council Grants (DP110104553) to M. E. and D. R. B., and (DP160104211) to D. R. B.

Commercial relationships: none.

Corresponding author: Reuben Rideaux.

Email: reuben.rideaux@anu.edu.au.

Address: Research School of Psychology, The Australian National University, Canberra, Australian Capital Territory, Australia.

## References

- Abeles, M., Bergman, H., Margalit, E., & Vaadia, E. (1993). Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *Journal of Neurophysiology*, 70(4), 1629–1638.
- Adelson, E. H., & Farid, H. (1999). Filtering reveals form in temporally structured displays. *Science*, 286(5448), 2231–2231.
- Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, 300, 523–525.
- Alais, D., Blake, R., & Lee, S. H. (1998). Visual features that vary together over time group together over space. *Nature Neuroscience*, 1(2), 160–164.
- Amano, K., Edwards, M., Badcock, D. R., & Nishida, S. Y. (2009). Adaptive pooling of visual motion signals by the human visual system revealed with a novel multi-element stimulus. *Journal of Vision*, 9(3):4, 1–25, doi:10.1167/9.3.4. [PubMed] [Article]
- Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, 40(1), 309–328.
- Bex, P. J., Simmers, A. J., & Dakin, S. C. (2001). Snakes and ladders: The role of temporal modulation in visual contour integration. *Vision Research*, 41(27), 3775–3782.
- Blake, R., & Lee, S. H. (2005). The role of temporal structure in human vision. *Behavioral and Cognitive Neuroscience Reviews*, 4(1), 21–42.
- Brainard, D. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436. Retrieved from <http://bbs.bioguider.com/images/upfile/2006-4/200641014348.pdf>
- Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience*, 10(6), 1157–1169.
- Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. In *Seminars in the Neurosciences*, (Vol. 2, pp. 263–275). Philadelphia, PA: Saunders Scientific Publications.
- Dickinson, J. E., Broderick, C., & Badcock, D. R. (2009). Selective attention contributes to global processing in vision. *Journal of Vision*, 9(2):6, 1–8, doi:10.1167/9.2.6. [PubMed] [Article]
- Diesmann, M., Gewaltig, M. O., & Aertsen, A. (1999). Stable propagation of synchronous spiking in cortical neural networks. *Nature*, 402(6761), 529–533.
- Driver, J., & Baylis, G. C. (1996). Edge-assignment and figure-ground segmentation in short-term visual matching. *Cognitive Psychology*, 31(3), 248–306, doi:10.1006/cogp.1996.0018.
- Edwards, M. (2009). Common-fate motion processing:

- Interaction of the On and Off pathways. *Vision Research*, 49(4), 429–438.
- Edwards, M., & Badcock, D. R. (1998). Discrimination of global-motion signal strength. *Vision Research*, 38(20), 3051–3056, doi:10.1016/S0042-6989(98)00018-2.
- Efron, R. (1970). The minimum duration of a perception. *Neuropsychologia*, 8(1), 57–63.
- Efron, R., & Lee, D. N. (1971). The visual persistence of a moving stroboscopically illuminated object. *The American Journal of Psychology*, 84(3), 365–375, doi:10.2307/1420468.
- Farid, H. (2002). Temporal synchrony in perceptual grouping: A critique. *Trends in Cognitive Sciences*, 6(7), 284–288.
- Farid, H., & Adelson, E. H. (2001). Synchrony does not promote grouping in temporally structured displays. *Nature Neuroscience*, 4(9), 875–876.
- Forte, J., Hogben, J. H., & Ross, J. (1999). Spatial limitations of temporal segmentation. *Vision Research*, 39(24), 4052–4061.
- Goodbourn, P. T., & Forte, J. D. (2013). Spatial limitations of fast temporal segmentation are best modeled by V1 receptive fields. *Journal of Vision*, 13(13):23, 1–18, doi:10.1167/13.13.23. [PubMed] [Article]
- Greenwood, J. A., & Edwards, M. (2009). The detection of multiple global directions: Capacity limits with spatially segregated and transparent-motion signals. *Journal of Vision*, 9(1):40, 1–15, doi:10.1167/9.1.40. [PubMed] [Article]
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1), 106–154, doi:10.1113/jphysiol.1962.sp006837.
- Hubel, D. H., & Wiesel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 28(2), 229–289. Retrieved from: <http://jn.physiology.org/content/jn/28/2/229.full.pdf>
- Kandil, F. I., & Fahle, M. (2004). Figure-ground segregation can rely on differences in motion direction. *Vision Research*, 44(27), 3177–3182.
- Kiper, D. C., Gegenfurtner, K. R., & Movshon, J. A. (1996). Cortical oscillatory responses do not affect visual segmentation. *Vision Research*, 36(4), 539–544.
- Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, 39(16), 2729–2737.
- Kösem, A., & van Wassenhove, V. (2012). Temporal structure in audiovisual sensory selection. *PloS One*, 7(7), e40936.
- Kovács, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: Effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Sciences, USA*, 90(16), 7495–7497.
- Lee, S. H., & Blake, R. (1999). Visual form created solely from temporal structure. *Science*, 284(5417), 1165–1168.
- Lee, S. H., & Blake, R. (2001). Neural synergy in visual grouping: When good continuation meets common fate. *Vision Research*, 41(16), 2057–2064.
- Leonards, U., Singer, W., & Fahle, M. (1996). The influence of temporal phase differences on texture segmentation. *Vision Research*, 36(17), 2689–2697.
- Lorceau, J., & Lalanne, C. (2008). Superposition catastrophe and form-motion binding. *Journal of Vision*, 8(8):13, 1–14, doi:10.1167/8.8.13. [PubMed] [Article]
- Milner, P. M. (1974). A model for visual shape recognition. *Psychological Review*, 81(6), 521–535. Retrieved from <http://dx.doi.org/10.1037/h0037149>
- Morley, D., & Badcock, D. R. (2016). Spatial integration of global motion. Manuscript in preparation.
- Morgan, M., & Castet, E. (2002). High temporal frequency synchrony is insufficient for perceptual grouping. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1490), 513–516.
- Motoyoshi, I. (2004). The role of spatial interactions in perceptual synchrony. *Journal of Vision*, 4(5):1, 352–361, doi:10.1167/4.5.1. [PubMed] [Article]
- Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, 18(1), 55–68, doi:10.1068/p180055.
- Newsome, W. T., & Pare, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *The Journal of Neuroscience*, 8(6), 2201–2211.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442, doi:10.1163/156856897X00366.
- Prins, N. (2013). The psi-marginal adaptive method: How to give nuisance parameters the attention they deserve (no more, no less). *Journal of Vision*, 13(7):3, 1–13, doi:10.1167/13.7.3. [PubMed] [Article]
- Prins, N., & Kingdom, F. A. A. (2009). Palamedes: Matlab routines for analyzing psychophysical data. Retrieved from <http://www.palamedestoolbox.org>



- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74(11), 1–29.
- Tang, M. F., Dickinson, J. E., Visser, T. A., Edwards, M., & Badcock, D. R. (2015). Role of form information in motion pooling and segmentation. *Journal of Vision*, 15(15):19, 1–18, doi:10.1167/15.15.19. [PubMed] [Article]
- Usher, M., & Donnelly, N. (1998). Visual synchrony affects binding and segmentation in perception. *Nature*, 394(6689), 179–182.
- von der Malsburg, C. (1994). The correlation theory of brain function. In E. Domany, J. L. van Hemmen, & K. Schulten (Eds.), *Models of neural networks II* (pp. 95–119). Berlin: Springer. Retrieved from [http://link.springer.com/chapter/10.1007/978-1-4612-4320-5\\_2](http://link.springer.com/chapter/10.1007/978-1-4612-4320-5_2). (Original work published 1981)